

Explaining Spatio-Temporal Dynamics in Carsharing: A Case Study of Amsterdam

Emergent Research Forum Paper

Konstantin Klemmer
University of Freiburg
konstantin.klemmer@is.uni-freiburg.de

Sebastian Wagner
University of Freiburg &
Geospin GmbH
swagner@geospin.de

Christoph Willing
University of Freiburg
christoph.willing@is.uni-freiburg.de

Tobias Brandt
University of Freiburg
tobias.brandt@is.uni-freiburg.de

Abstract

We investigate customer mobility behavior by examining free-floating carsharing demand dynamics. For this purpose, we analyze rental data of a major carsharing provider in the city of Amsterdam in combination with points of interest (POIs). Connecting POI data to carsharing trips and stratifying the data along 6-hour intervals allows us to illustrate the spatio-temporal dimensions of carsharing usage, i.e. how carsharing demand changes over time and how it shifts spatially within the provider's business area. We cluster the point data using kernel density estimation and apply a generalized linear model with Gamma distributed values on the sampled data. Our results indicate that, depending on the hour of the day, different POI categories have different, yet significant, impact on trip destinations. Our insights advance the understanding of when and for what purpose customers use carsharing, enabling providers to predict demand in existing and new business areas.

Keywords

Carsharing, Spatial Analytics, Location-based Services, Visualization

Introduction

During the last decade, carsharing has emerged as one of the most prominent examples of the sharing economy. In urban areas with a well-developed carsharing infrastructure, shared vehicles can be even more attractive than a personal car. By using carsharing, customers can save taxes and maintenance costs, while still being able to access vehicles whenever needed. The elimination of the fixed costs of car ownership also enables customers who previously could not afford these expenses to enjoy the benefits of private car travel (Martin et al. 2010). Furthermore, the increasing adoption of carsharing results in a reduction of CO₂ emissions, which contributes to solving the pollution problem of many metropolitan areas (Firnborn and Müller 2011). In addition, several carsharing companies offer, either partly or exclusively, hybrid and electric vehicles and thus contribute to environmental sustainability even further.

Today's carsharing systems face complex challenges. An advanced method of carsharing is the free-floating approach: In this model, vehicles may be driven to any location within a predefined operational area. Compared to a model with fixed parking spots or stations, free-floating carsharing offers customers more flexibility and enables new transportation opportunities such as one-way trips or the combination of carsharing and public transport. However, the problem of vehicle relocation remains unresolved, since within the operating area there are regions of relatively high demand (underflow region) and relatively low demand (overflow region), which exhibit fewer and more vehicles than needed, respectively (Lee and Park 2014). The essential step in resolving this issue is to figure out why these regions emerge in the first place. Therefore, the determinants of vehicle distribution need to be identified and the behavior of carsharing users has to be analyzed, which is why in this paper we aim to answer the following questions:

1) Why and when do customers use carsharing?

Are the different purposes of the trips somehow connected to specific time periods? We aim to find the sources contributing to demand imbalances and examine, whether these unbalanced regions are temporarily independent or change during the course of a given time period.

2) How can these insights be used to address the *vehicle relocation problem*?

With the gathered information regarding consumer behavior, we can add to the development of reliable demand prediction. Furthermore, our results offer new insight into the spatio-temporal imbalances of shared vehicle utilization, which is a crucial factor in understanding relocation issues.

Data sources

Since our investigation applies to a free-floating environment, it is important to fully understand the functional principles of such a carsharing system. To this end, we use data of a major carsharing provider in the city of Amsterdam. All relevant observations are gathered from within the operational area, which is illustrated by the blue lines in Figure 1. The vehicle data was collected between February 2015 and April 2015 and includes over 280,000 individual trips undertaken by a customer base of more than 10,000 members. Every trip has a specific end point which consequently is also the starting point of the next trip with the same vehicle.

In order to investigate temporal dynamics of vehicle distribution, we need to split our data into time intervals. Initially, we assign every trip to the subsets ‘Workday’ (Monday to Friday), ‘Saturday’ and ‘Sunday’. The sum V of our trip ending point data is therefore given as:

$$V = \{V_{Mo-Fr}, V_{Sa}, V_{Su}\} \quad (\text{Eq. 1})$$

The above subsets are then further divided into time intervals of six hours, to represent nighttime (0-6h), morning (6-12h), afternoon (12-18h) and evening hours (18-24h).

$$V_n = \{V_n^{0-6}, V_n^{6-12}, V_n^{12-18}, V_n^{18-24}\} \quad (\text{Eq. 2})$$

We thus obtain 12 subsets: three day-of-week samples, each consisting of four time-of-day samples.

These subsets enable us to investigate and visualize the intraday and intraweek imbalances in vehicle locations. Figure 2 shows heat maps of trip ending points from our workday sample. The four heat maps

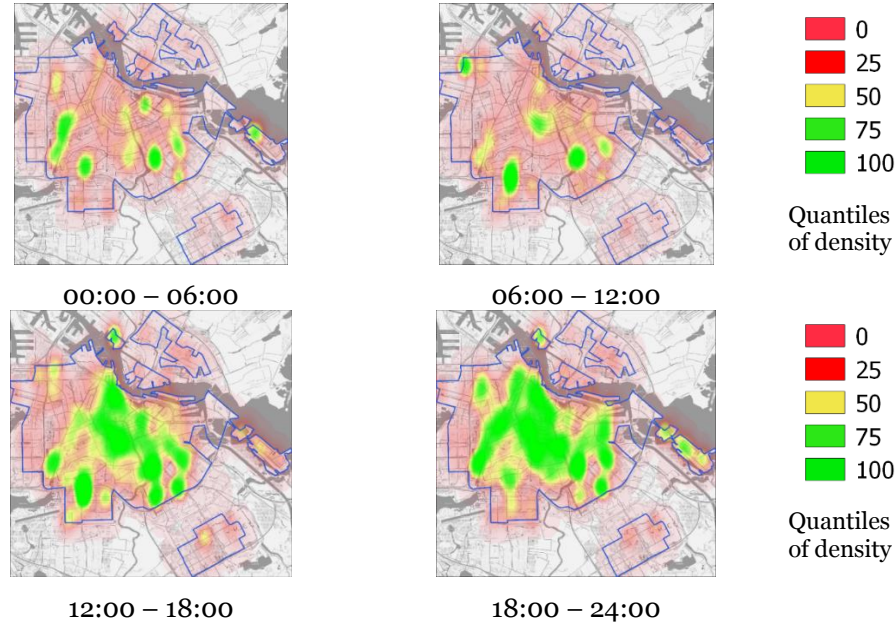


Figure 1. Density of trip ending points on workdays (Mo - Fr)

each refer to one of the four predefined time-of-day intervals. Red areas contain a low number of trip endpoints, while green areas imply a high density of trip ending points.

It also becomes apparent that different areas may be unequally attractive for carsharing during different times of the day. The graphical analysis thus supports our decision to cluster our data, as trip patterns change across the different samples.

In order to explain the observed distribution imbalances and spatial shifts, we need additional data. For this, we use point-of-interest (POI) data provided by a large open-access geomapping service. We extract more than 60,000 data points, which are arranged into 91 different categories such as cafes, banks or bus stations. Following Bendler et al. (2014) we cluster POI categories and create eight supersets: *Retail and Services*, *Gastronomy*, *Social and Religious*, *Public Authority*, *Finance*, *Transportation*, *Leisure Time*, *Party*. Each POI data point is then assigned to one of these eight supersets, and the set of all POIs can be expressed as:

$$I = \{I_{R\&S}, I_G, I_{S\&R}, I_{PA}, I_F, I_T, I_{LT}, I_P\} \quad (\text{Eq. 3})$$

Figure 2 shows heat maps of POI locations for the two supersets ‘Finance’ and ‘Transportation’. The samples are comparable in size and therefore well suited to demonstrate the spatial variations between different clusters.

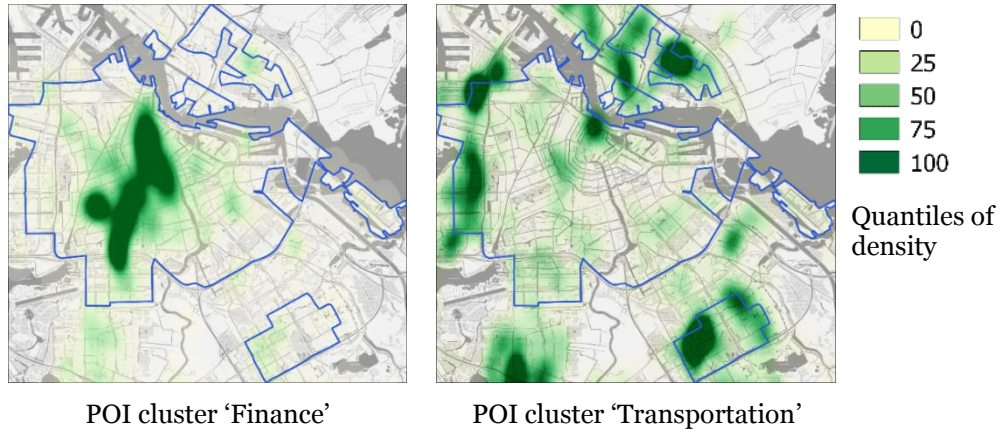


Figure 2. Density of POIs from selected clusters

Approach

In combination, the vehicle and POI datasets allow us to explain carsharing activity by the proximity of rental endpoints to POIs from the clustered supersets. In order to merge the two datasets we apply a kernel density estimation (KDE).

We generate a point-grid within the operating area, containing 6,664 evenly distributed points which serve as kernels. Each kernel is then assigned all POIs and rental destinations within a predefined bandwidth. The number of points assigned to the kernel as well as the distance to these points defines the kernel's density value. Assigning points to kernels and working with densities allows us to drastically reduce zero values which would otherwise be overwhelming in spatial data (Nakaya and Yano 2010, Hallin et al. 2004). Below, we illustrate the application of KDE to our dataset. First, we create the grid of regularly distributed points within the operational area. Second, we add the point data to be sampled, i.e. trip ending points and POIs. Finally, we compute the kernel density value representing the sampled points for each kernel. We repeat this step for the 12 trip ending point samples V and the eight POI samples I and obtain 20 ($V + I$) kernel density values for all 6,664 kernels.

Next, we want to investigate how the density of the different POI groups (explanatory variables) influence the density of trip ending points (dependent variable). A Gamma type distribution appears to be most suited to our data, as it accounts for continuous numbers (Stacy and Mihram 1965). We therefore apply a generalized linear model (GLM) with Gamma distributed values to every trip ending point dataset.

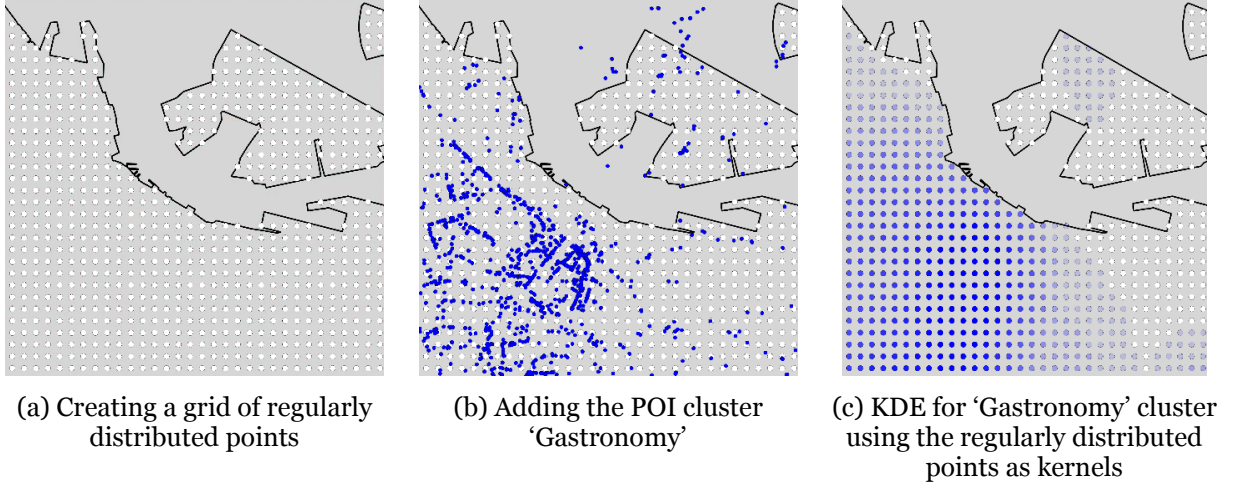


Figure 3. Kernel density estimation in spatial context

Empirical Results

As can be observed by our regression results in Table 1, the 'Leisure Time' cluster, including gyms, museums and movie theaters, shows a significantly negative influence during work hours while being significantly positive on weekday evenings and until Saturday noon. As expected, 'Party' has a large and highly significant coefficient on Saturday evenings and less pronounced or negative values otherwise. 'Public Authority' POIs have a significantly positive influence only during weekday morning hours. The results, however, have to be regarded with caution, as the cluster exhibits a high variance inflation factor (VIF). 'Social & Religious', including schools, libraries and churches, also shows a slightly positive influence during weekday working hours, but has a significant negative impact during the weekend, except for Sunday mornings. Presumably, in this timeslot the negative influence of schools and libraries, which are closed then, is offset by the positive influence of churches. The 'Gastronomy' results appear counterintuitive, e.g. negative impact on Saturday evening, but this cluster also exhibits the highest VIF and thus lacks explanatory value. With the exception of Sunday morning, 'Transportation' is the only cluster, with significant positive estimates across all time-slots. This can be explained by the complementary nature of carsharing and public transport. The coefficients for 'Finance' are mostly significant but difficult to interpret and require additional investigation. The same is true for 'Retail & Services'. This is by far the largest cluster and therefore also exhibits a high VIF value.

Table 1. Results of generalized linear regression

	Monday-Friday				Saturday				Sunday			
	00-06h	06-12h	12-18h	18-24h	00-06h	06-12h	12-18h	18-24h	00-06h	06-12h	12-18h	18-24h
(Intercept)	2.69***	1.75***	1.39***	1.98***	2.18***	3.13***	1.21***	1.09***	2.17***	3.41***	1.35***	1.15***
Leisure Time	0.03***	-0.01***	0.00*	0.00***	0.14***	0.10***	0.01	-0.05***	-0.01	0.00	0.03**	0.01
Party	-0.17***	0.02	0.03***	-0.03***	-1.06***	-0.83***	0.06	0.26***	-0.08	0.48*	0.02	0.10*
Publ. Auth.	-0.17***	0.12**	0.02	-0.10***	-0.05	-0.52	-0.37***	-0.52***	-0.33	-0.21	-0.49***	-0.75***
Soc. & Rel.	0.00	0.01**	0.00	-0.01***	-0.09***	-0.12***	-0.04***	-0.05***	-0.05***	-0.02	-0.03***	-0.01**
Gastro.	0.01**	-0.01***	-0.01***	0.00	0.04***	0.04**	-0.02***	-0.03***	-0.01	-0.16***	-0.04***	-0.02***
Transp.	0.27***	0.05***	0.06***	0.10***	0.59***	0.50***	0.18***	0.32***	0.37***	0.21	0.26***	0.75***
Finance	0.00	-0.08***	-0.03***	0.01***	0.18***	0.14**	0.00	0.15***	0.28***	-0.18**	-0.07***	0.10***
Ret. & Serv.	-0.01***	0.01***	0.00**	0.00***	-0.06***	-0.06***	-0.01***	-0.01**	-0.05***	0.00	-0.01*	-0.03***
Pseudo R ²	0.21	0.18	0.24	0.42	0.30	0.33	0.40	0.44	0.29	0.42	0.39	0.36
Significance level at *** 0.001, ** 0.01, * 0.05, n=6664 observations												

Conclusion

Carsharing is growing rapidly in metropolitan areas all over the world and the need for its efficiency and user-friendliness is constantly increasing. Information Systems have emerged as a powerful instrument to connect providers and customers, which also applies to carsharing, whose customers are offered a maximum degree of flexibility. Flexibility, however, comes at the price of growing complexity and optimization challenges. In our research we therefore investigate carsharing customer behavior by connecting trip activity with POIs and presents the results in a spatio-temporal context. We use a vast dataset of a major carsharing provider in Amsterdam and examine how trip destinations are connected to different POI categories. By using kernel density estimation and applying a generalized linear model with Gamma distributed values, we can show that different POI types have a different, yet significant, impact on trip destinations. We thus confirm the results of a recent study by Wagner et al. (2016) and extend it by providing new insight into the temporal sustainability of POI influence. We observe noticeable changes in sign and significance across the eight POI clusters. Consequently, we infer that POI influence on carsharing destinations depends on the respective time period. Our work contributes to current research in two ways:

1) We assist in customer understanding by investigating *when* and *for what purpose* carsharing is used. Our findings suggest that in this connection, the temporal and spatial dimensions are not independent. Rather, different POI categories have a different influence on trip destinations depending on the time of day and the day of the week. We draw the conclusion that the spatio-temporal dynamics of carsharing utilization can, to some extent, be explained by POIs. This knowledge not only enables managers to understand their customers' behavior, but also equips them with a valuable tool for demand prediction, which directly leads us to the second research question.

2) By providing new insight into shared vehicle demand, we can extract useful information about the development of underflow/overflow regions and hence add to the resolution of the relocation problem. Since the first step in solving any problem is to understand all possible angles, we hereby extend the common knowledge and put it into a spatio-temporal perspective. Our results indicate that unbalanced regions are not only determined by POIs, but also shift spatially, which we visualized using GIS techniques. Carsharing businesses may thus use our findings within their relocation operations or even develop improved approaches.

The sharing economy in general and carsharing in particular are still in the early stages of their evolution with a lot of room left for improvement and growth. Their development is inseparably bound to the development of information systems. The essential challenge for IS developers is to convert an immensely complex mathematical process to a user-friendly, comprehensible tool for everyday use.

References

- Bendler, J., Ratku, A., and Neumann, D. 2014. "Crime Mapping through Geo-Spatial Social Media Activity," *Thirty Fifth International Conference on Information Systems, Auckland* (2014).
- Firnkorn, J., and Müller, M. 2011. "What will be the environmental effects of new free floating car-sharing systems? The case of car2go in Ulm," *Ecological Economics* (70:8), pp. 1519–1528.
- Hallin, M., Lu, Z., and Tran, L. T. 2004. "Kernel density estimation for spatial processes: the L1 theory," *Journal of Multivariate Analysis* (88:1), pp. 61–75.
- Lee, J., and Park, G. L. 2014. "Cluster-based Vehicle Redistribution scheme based on Genetic Algorithms for Electric Vehicle Sharing Systems," *International Journal of Software Engineering & Its Applications* (8:9), pp. 147–158.
- Martin, E., Shaheen, S. A., and Lidicker, J. 2010. "Impact of carsharing on household vehicle holdings," *Transportation Research Record: Journal of the Transportation Research Board* (2143:1), pp. 150–158.
- Nakaya, T., and Yano, K. 2010. "Visualising Crime Clusters in a Space-time Cube: An Exploratory Data-analysis Approach Using Space-time Kernel Density Estimation and Scan Statistics," *Transactions in GIS* (14:3), pp. 223–239.
- Stacy, E. W., and Mihram, G. A. 1965. "Parameter Estimation for a Generalized Gamma Distribution," *Technometrics* (7:3), p. 349.
- Wagner, S., Brandt, T., and Neumann, D. 2016. "In Free Float: Developing Business Analytics Support for Carsharing Providers," *Omega* (59), pp. 4–14.